**Analyzing Dark Side of Social media Using Qlik**

**Introduction:**

ABC Corporation, recognized for its commitment to digital wellness, is launching a project to address and mitigate the impact of time-wasting activities on social media through advanced data analytics and visualization techniques using Qlik. The project centers on analyzing a detailed dataset that includes factors such as time spent on various social media platforms, user engagement levels, content type, and interaction patterns. By leveraging this data, the company aims to identify the key drivers behind time-wasting behaviors and develop strategies to enhance user productivity and well-being.

The analysis will employ a decomposition tree to dissect time-wasting activities by different variables, offering a clear view of how each factor contributes to overall social media usage patterns. Additionally, the project will involve the creation of several calculated columns and measures to deepen the dataset's analytical insights. Visualizations such as heat maps, pie charts, line charts, and bubble charts will be used to effectively communicate the findings. By implementing this solution, ABC Corporation seeks to help users manage their social media time more effectively, promote healthier digital habits, and support overall productivity, thereby reinforcing its position as a leader in promoting digital wellness and responsible technology use.

**Scenario 1:**

ABC has been experiencing a noticeable decline in productivity among its remote teams. Initial investigations suggest that excessive social media use might be contributing to this issue. To address the problem, ABC decided to analyze social media usage patterns using Qlik. The analysis focused on employees' time spent on various social media platforms, engagement levels, and work output metrics. By utilizing a decomposition tree to break down this data by team, job role, and time of day, ABC discovered that certain teams were more prone to social media distractions, particularly during peak work hours. In response, the company implemented targeted interventions such as time management workshops and productivity tools.

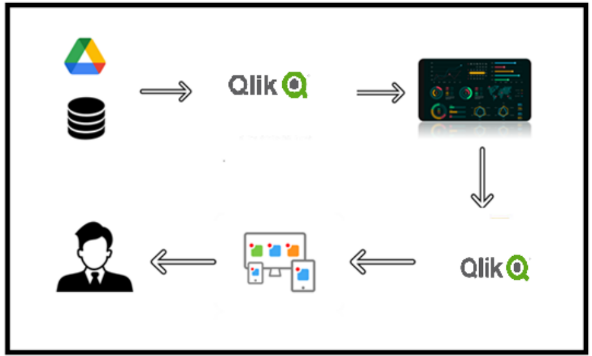
**Scenario 2:**

ABC High School has observed a troubling decline in academic performance among its students. Early findings suggest that excessive social media use might be a significant factor. To investigate, the school utilized Qlik to analyze data on students' social media usage patterns, including time spent on different platforms, types of content engaged with, and academic performance metrics. The decomposition tree was used to segment the data by grade level, subject area, and study habits. The analysis revealed a strong correlation between high social media usage and lower academic performance, especially among students in their final year. In response, ABC High School developed targeted educational programs and workshops on digital wellness and effective study techniques.

**Scenario 3:**

ABC Foundation has been concerned about increasing mental health issues among young adults, suspecting that social media might be playing a role. To explore this, the foundation employed Qlik to analyze data on social media usage, including time spent on various platforms, types of content engaged with, and self-reported mental health symptoms. The decomposition tree helped break down this data by age group, social media platform, and usage frequency. The analysis revealed a significant link between high social media engagement, particularly on platforms promoting unrealistic standards, and increased anxiety and depression. Using these insights, ABC Foundation is developing targeted mental health awareness campaigns and digital wellness programs. These efforts aim to promote healthier social media habits and provide necessary support resources, improving the mental well-being of young adults in the community.

**Technical Architecture:**

****

**Project Flow**

To accomplish this, we have to complete all the activities listed below,

* Data Collection
  + Collect the dataset,
* Connect Data with Qlik
* Data Preparation
  + Prepare the Data for Visualization
* Data Visualizations
  + Visualizations
* Dashboard
  + Responsive and Design of Dashboard
* Report
* Report Creation
* Performance Testing
  + Utilization of Data Filters
  + No. of Calculation fields
  + No. of Visualizations/Graphs
* Project Demonstration & Documentation
  + Record explanation Video for project end to end solution
  + Project Documentation-Step by step project development procedure

**Milestone 1: Data Collection & Extraction from Database**

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, evaluate outcomes and generate insights from the data.

**Activity 1: Downloading the dataset**

Please use the link to download the dataset: [Click Here](https://www.kaggle.com/datasets/muhammadroshaanriaz/time-wasters-on-social-media)

**Activity 1.1: Understand the data**

Data contains all the meta information regarding the columns described in the CSV files

**Column Description of the Dataset:**

* **User ID:** A unique identifier assigned to each user.
* **Age:** The age of the user.
* **Gender:** The gender of the user.
* **Location:** The geographical location of the user.
* **Income:** The annual income of the user.
* **Debt:** Tells If the is in Debt or Not.
* **Owns Property:** Indicates whether the user owns any property (Yes/No).
* **Profession:** The profession or job title of the user.
* **Demographics:** Additional demographic information about the user (Rural or Urban Life).
* **Platform:** The social media platform used by the user (e.g., Facebook, Instagram, Tik Tok).
* **Total Time Spent:** The total time the user has spent on the platform.
* **Number of Sessions:** The number of sessions the user has had on the platform.
* **Video ID:** A unique identifier for each video watched.
* **Video Category:** The category of the video watched (e.g., Entertainment, Gaming, Pranks, Vlog).
* **Video Length:** The length of the video watched.
* **Engagement:** The engagement level of the user with the video (e.g., Likes, Comments).
* **Importance Score:** A score representing the perceived importance of the video to the user.
* **Time Spent on Video:** The amount of time the user spent watching the video.
* **Number of Videos Watched:** The total number of videos watched by the user.
* **Scroll Rate:** The rate at which the user scrolls through content.
* **Frequency:** How frequently the user logs into the platform.
* **Productivity Loss:** The amount of productivity lost due to time spent on social media.
* **Satisfaction:** The satisfaction level of the user with the content consumed.
* **Watch Reason:** The reason why the user watched the video (e.g., Entertainment, Information).
* **Device Type:** The type of device used to access the platform (e.g., Mobile, Desktop).
* **OS:** The operating system of the device used.
* **Watch Time:** The specific time of day when the user watched the video.
* **Self-Control:** The user's self-assessed level of self-control while using the platform.
* **Addiction Level:** The user's self-assessed level of addiction to social media.
* **Current Activity:** The activity the user was engaged in before using the platform.
* **Connection Type:** The type of internet connection used by the user (e.g., Wi-Fi, Mobile Data).

**Activity 2: Prepare the Data for Visualization**

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. This process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency. Since the data is already cleaned, we can move to visualization.

3.1: Data Loading

[Click Here](https://drive.google.com/file/d/1CeQ5GR_AU_hv-mCak44pvuXreOmU2Yb-/view?usp=sharing)

3.2 Data Cleaning

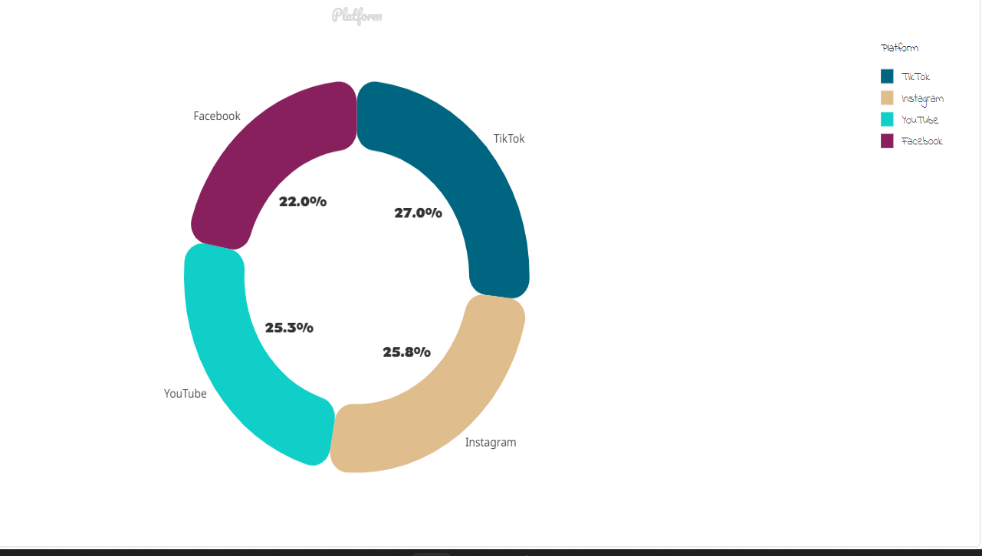
[Click Here](https://drive.google.com/file/d/1m8JB1VvjWquMsFRHkVzniL_oPURpgWBL/view?usp=sharing)

**Milestone 2: Data Visualization**

Data visualization is the process of creating graphical representations of data to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

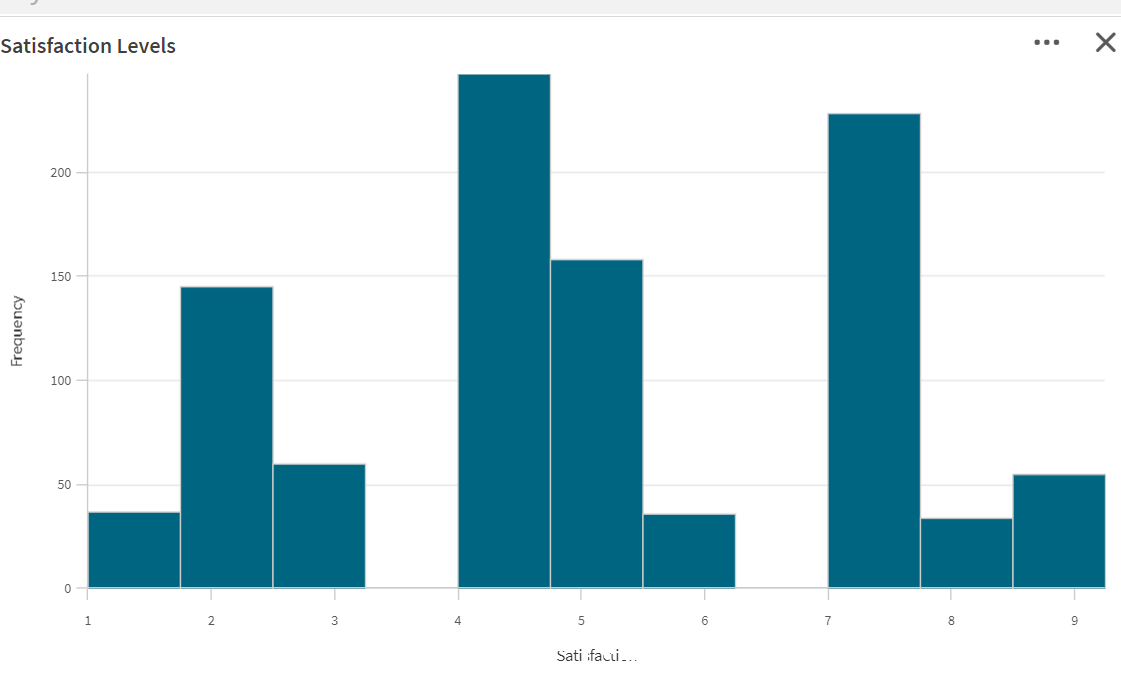
**Activity 1: Dark side of social Media**

**Activity 1.1: Platform Usage.**

****

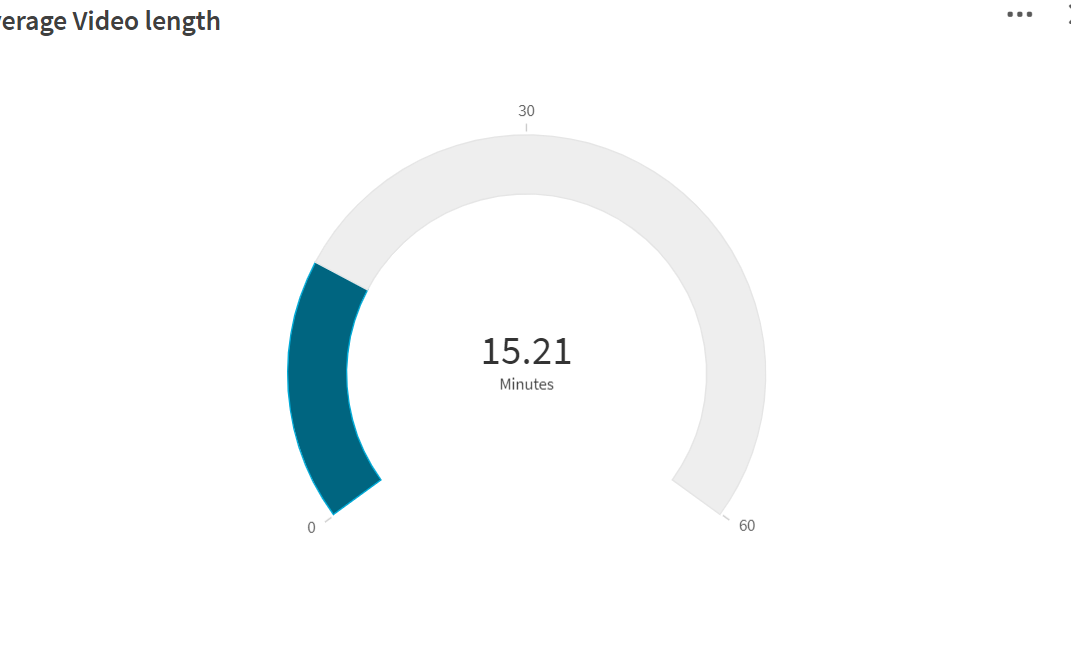
**Insight:** It reveals that Tik Tok dominates the social media landscape with a 27.0% share, closely followed by Instagram (25.8%) and YouTube (25.5%). Facebook maintains relevance at 22.0%, while other platforms collectively contribute a negligible 0%. This distribution sheds light on user engagement and market presence across these popular channels.

**Activity 1.2: Satisfaction Levels**

****

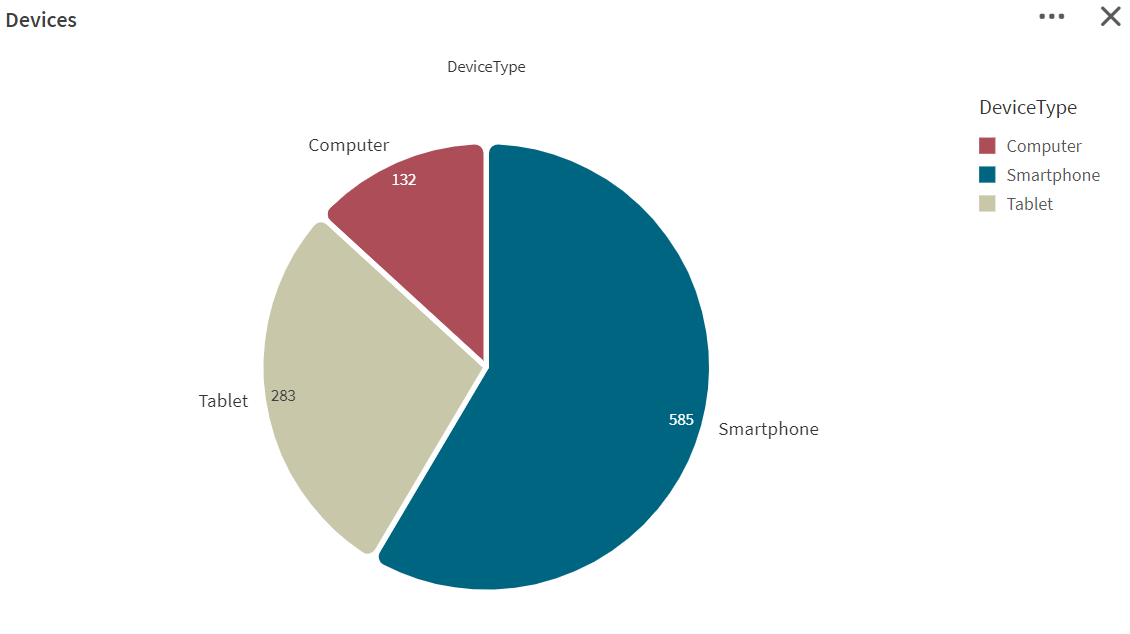
**Insight:** The majority rating their experience at level 8. This suggests that the product or service generally meets or exceeds expectations. However, there is a range of satisfaction levels, including some lower ratings (e.g., levels 4-6).

**Activity 1.3Average Video Length People Watching.**

****

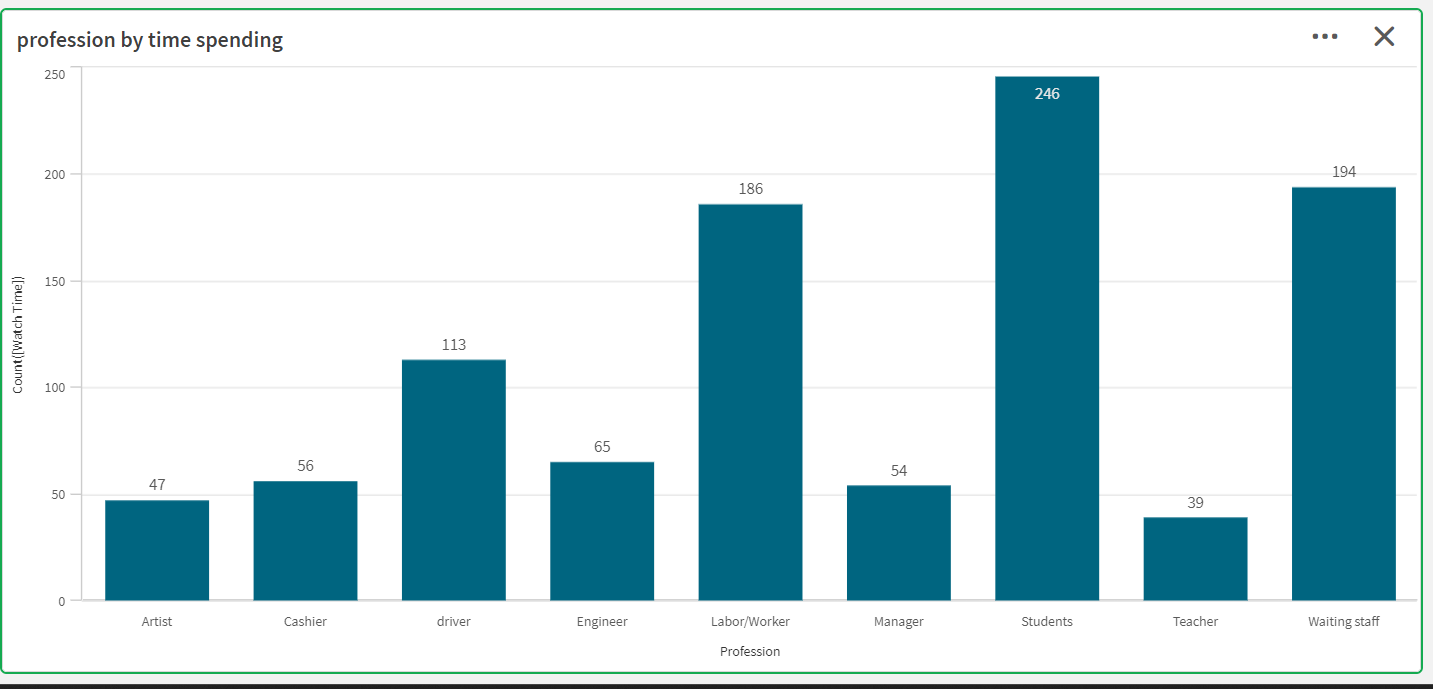
**Insight:** It shows the average number of minute that the people are watching the video. The average is approximately 15mins 21 sec.

**Activity 1.4: Types of Devices**

****

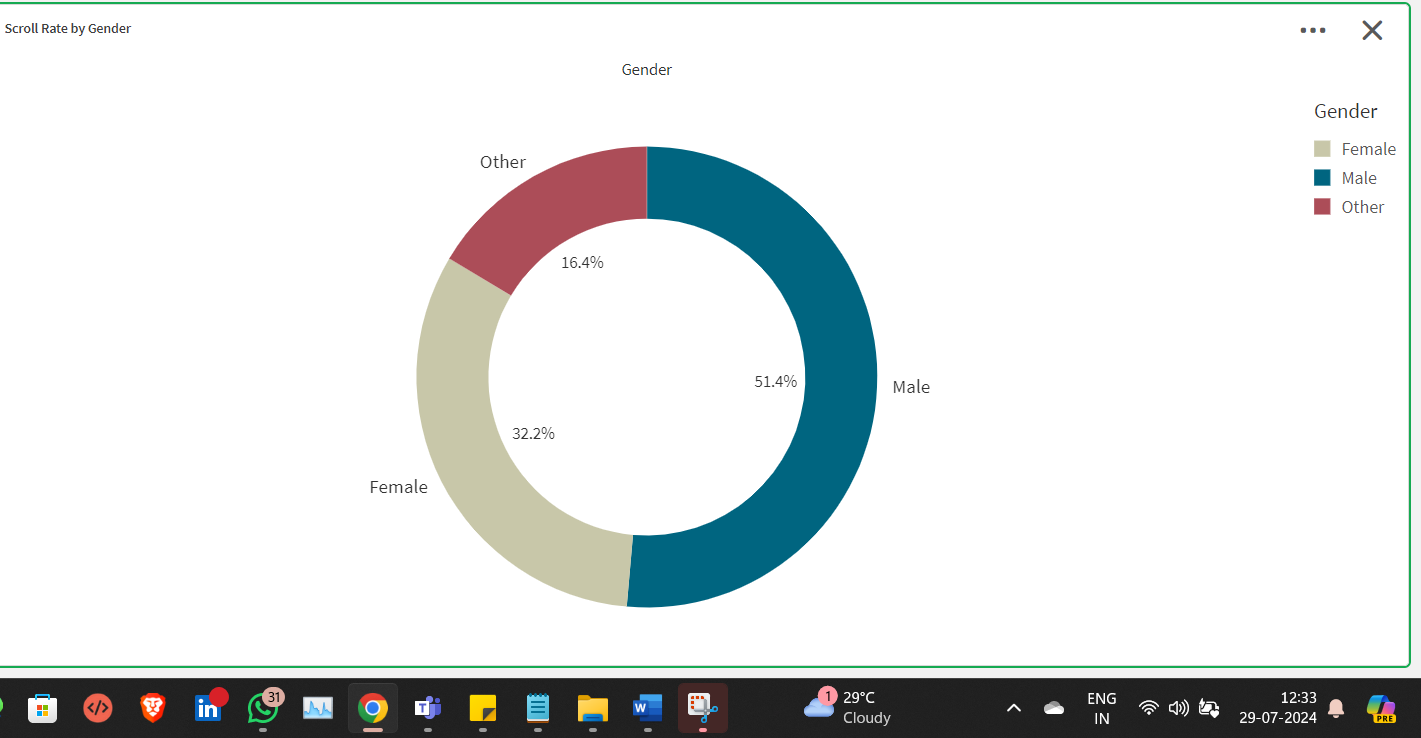
**Insight** The pie chart shows that smartphones are the most common device type, with 585 instances, followed by tablets at 283 occurrences, and computers at 132. This distribution underscores the dominance of smartphones in the dataset.

**Activity 1.5: Time spending by Professions.**



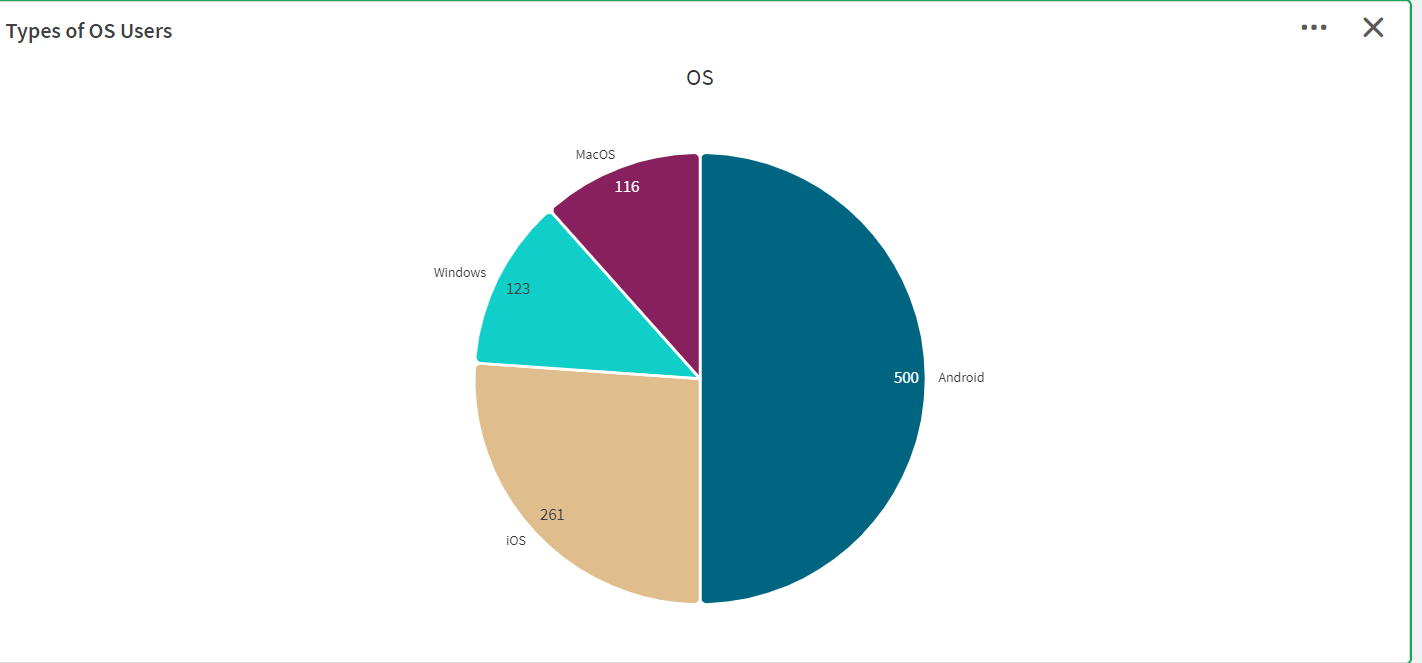
**Insights: “**Profession by Time Spending” reveals that scientists work the most, with about 246 hours per week. Laborers follow at 156 hours, while teachers spend around 194 hours. Managers work about 54 hours, and artists and cashiers have varying commitments, with artists at 47 hours and cashiers at 56 hours.

**Activity 1.6: Scroll Rate by Gender**

****

**Insights:** The gender distribution shows that males make up the largest segment at 51.8%, followed by females at 32.2%, and others at 14.6%. This breakdown offers a clear overview of the gender composition within the analyzed group, aiding in demographic analysis and decision-making.

**Activity 1.7: Types of OS Users.**

****

**Insights :** As of February 2024, Windows leads the desktop and tablet OS market with a 68.15% share, while Apple's iOS holds 28.39% of the mobile market, and macOS has 5.72% for desktops. Android dominates smartphones with a 70.87% share, and Apple’s iPad leads tablets with 55.53%, followed by Android at 43.97%.

**NOTE:** Video Explanations for the above Visualizations are in Dashboard and Report sections.

**Milestone 4: Dashboard**

A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

**Activity 1- Responsive and Design of Dashboard**

**Explanation video link:**

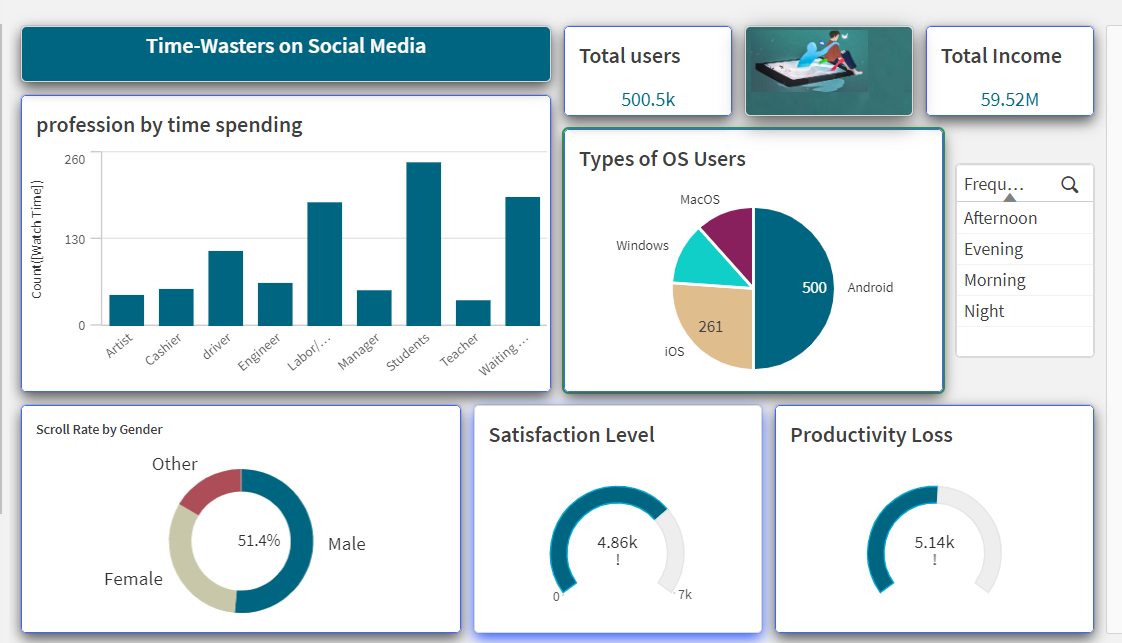
[Click Here](https://drive.google.com/file/d/1imY4iJsR14vJu_5fSsu7xTDWsPxTaqbT/view?usp=sharing)

**Dashboard1:**

****

**Insights:** The infographic provides insights into social media usage patterns. Facebook and Tik Tok each capture 27% of users' time, while YouTube holds 24% and Instagram 22%. These platforms cater to varied interests and offer diverse content. User satisfaction levels vary, highlighting the need for targeted improvements and content strategies. Although demographic details are cropped, they likely shed light on age groups and gender distribution, essential for effective marketing. The line graph on video time indicates fluctuations, suggesting optimal times for content scheduling and ad placements. Lastly, the pie chart shows smartphones are the preferred devices, emphasizing the importance of mobile-friendly content.

**Dashboard2:**

****

**Insights:** The infographic highlights social media usage across different professions, revealing that artists and students spend the most time on these platforms, while managers spend the least. The pie chart shows operating system preferences among users, with Windows leading, followed by Android and MacOS. Gender differences in scrolling behavior are evident, as females scroll slightly more frequently than males. Additionally, satisfaction levels are moderate, but there is noticeable productivity loss associated with social media usage.

**Milestone 5: Storytelling**

A report is a comprehensive document that provides a detailed and structured account of data analysis, findings, and insights. It is typically used for in-depth analysis, documentation, and communication of results. Reports are suitable for a diverse audience, including decision-makers, analysts, and stakeholders who need a comprehensive understanding of the data.

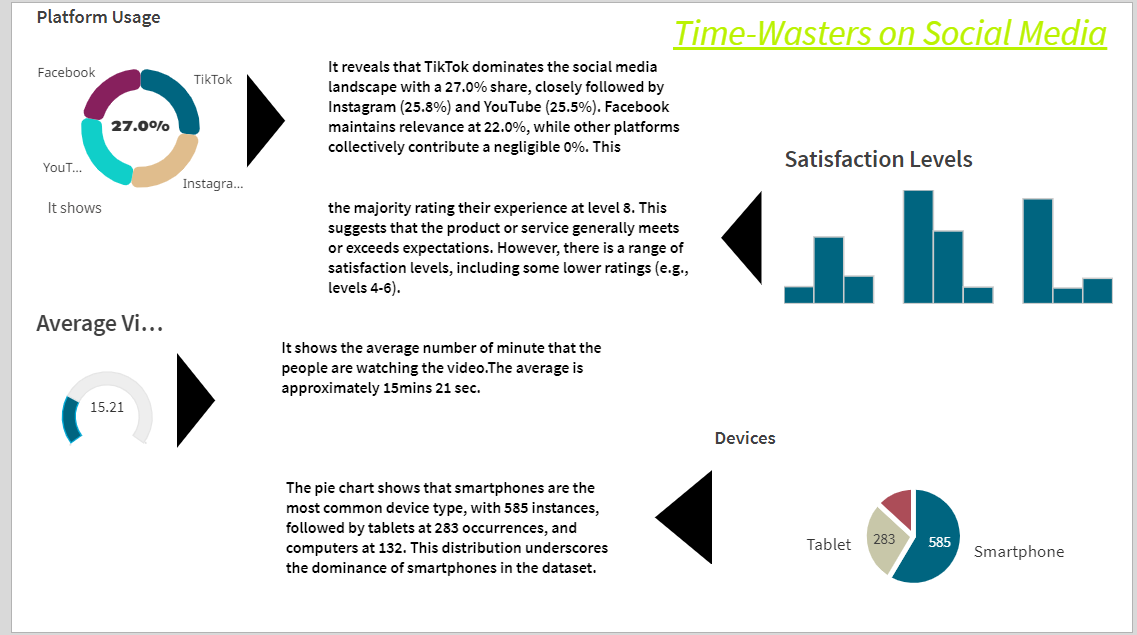
**Activity 1: Design of Storytelling**

Designing a report in Power BI involves connecting to data sources, creating visualizations like charts and graphs, customizing their appearance and interactivity, organizing them logically on the canvas, formatting elements for consistency and clarity, and optionally creating dashboards for a summarized view. Throughout the process, it's essential to consider the audience's needs and ensure the report effectively communicates insights from the data. Finally, iterate based on feedback to continually improve the report's design and usefulness.

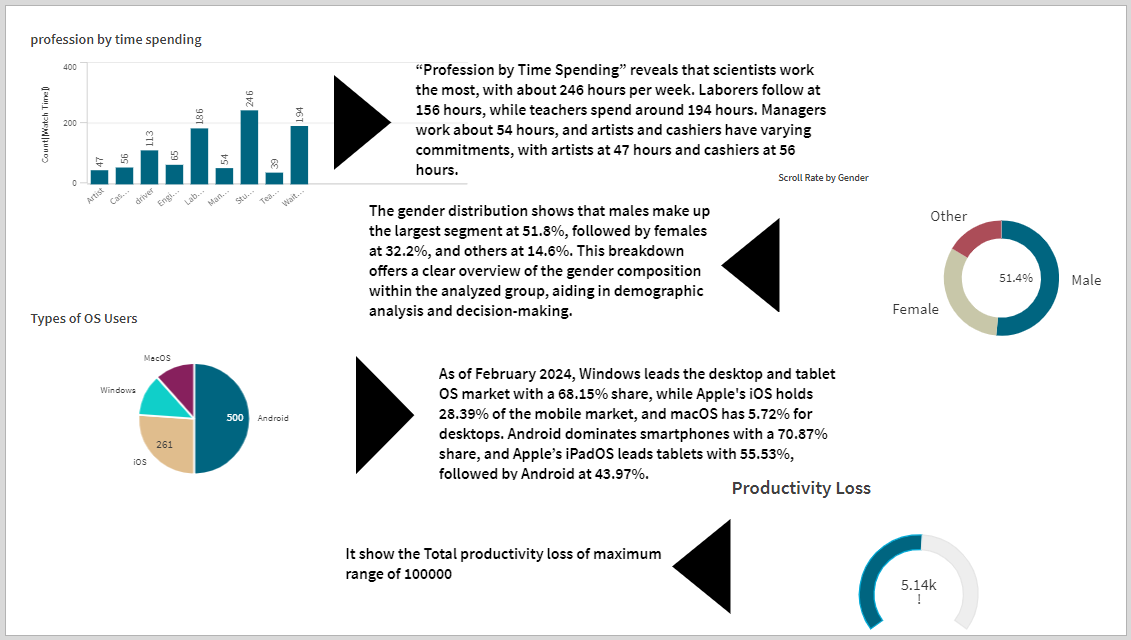
**Explanation video link:**

[Click Here](https://drive.google.com/file/d/1BKtVACT44_9c2g9_xJ0O7oFsTUZk-NFJ/view?usp=sharing)

**Report1:**

****

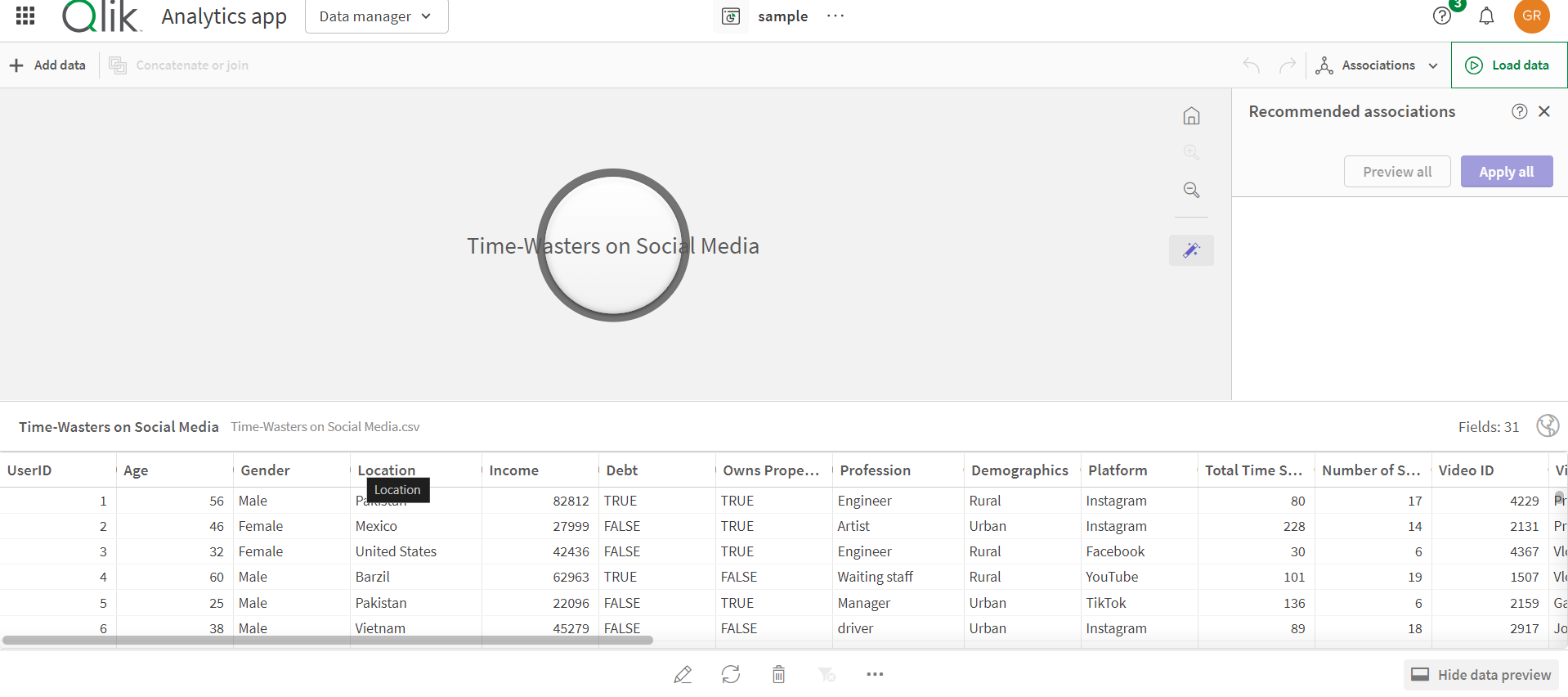
**Report 2:**

****

**Milestone 6: Performance Testing**

**Activity 1: Amount of Data Loaded**

"Amount of Data Loaded" refers to the quantity or volume of data that has been imported, retrieved, or loaded into a system, software application, database, or any other data storage or processing environment. It's a measure of how much data has been successfully processed and made available for analysis, manipulation, or use within the system.

****

**Activity 2.2: No of Visualizations/ Graphs**

Certainly! Here are the number of visualizations based on your description:

* **Time-Wasters on social media:** Chart showing time allocation by profession.
* **Types of OS Users**: Pie chart breaking down users by operating system preference.
* **Scroll Rate by Gender:** Metric indicating scrolling frequency by gender.
* **Satisfaction Level and Productivity Loss:** Semi-circle gauges representing satisfaction levels and productivity loss.

 **Platform Usage:** The chart shows that Facebook and TikTok each account for 27% of users' time, YouTube for 24%, and Instagram for 22%. These platforms cater to varied interests and offer diverse content.

 **Satisfaction Levels:** The bar graph displays varying user satisfaction levels, indicating different degrees of contentment across platforms. Understanding these levels can help in refining platform features and content strategies.

 **Demographics:** Although the demographic graph is cropped, it likely provides insights into user age groups, gender distribution, or other demographic details, essential for effective marketing.

 **Time Spent on Videos:** The line graph shows fluctuations in video consumption, suggesting peak times for content scheduling and ad placements.

 **Device Preference:** The pie chart reveals that smartphones are the most used devices for social media, with **365** units, highlighting the need for mobile-optimized content.

**Milestone 7: Project Demonstration & Documentation**

Below mentioned deliverables to be submitted along with other deliverables

**Activity 1: - Record explanation Video for the project's end-to-end solution**

**Activity 2: - Project Documentation-Step by step project development procedure**

Create document as per the template provided